PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/00 | A2

(11) International Publication Number: WO 99/00407

(43) International Publication Date: 7 January 1999 (07.01.99)

(21) International Application Number: PCT/EP98/04033

(22) International Filing Date: 25 June 1998 (25.06.98)

(30) Priority Data:
08/884,389 27 June 1997 (27.06.97) US

(71) Applicant *(for all designated States except US)*: PLANT GENETIC SYSTEMS N.V. [BE/BE]; Jozef Plateaustraat 22, B–9000 Gent (BE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: DE ROECK, Sandra [BE/BE]; Oudenhovestraat 15, B–9620 Zottegem (BE). VAN RIE, Jeroen [BE/BE]; Gravin Johannalaan 10, B–9900 Eeklo (BE).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: IMPROVED BACILLUS THURINGIENSIS TOXIN

(57) Abstract

New improved Cry9C proteins, having significantly increased toxicity, and DNA sequences encoding these proteins, were designed. Analysis of amino acid positions in domain II of the Cry9C protein by protein mutagenesis identified amino acids involved in insect toxicity. Random replacement of these amino acids identifies proteins with improved toxicity. A combination of identified improved amino acids in a single protein yields modified Cry9C proteins with significantly improved toxicity.

1, 3, 4,

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Paso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

IMPROVED BACILLUS THURINGIENSIS TOXIN

BACKGROUND OF THE INVENTION

(i) Field of the Invention

The present invention provides new improved proteins derived from a *Bacillus thuringiensis* Cry9C crystal protein. In accordance with this invention, amino acid positions in a Cry9C protein were identified as involved in insect toxicity. Further in accordance with this invention are provided modified Cry9C proteins with increased of decreased toxicity to an insect species, and DNA sequences encoding such modified Cry9C proteins. Plants can be protected from insect damage by expressing a chimeric gene encoding an improved Cry9C protein with an increased toxicity to an insect species.

(ii) Description of Related Art

Bacillus thuringiensis (Bt)-derived proteins are currently widely used to protect plants from insects by expression of such proteins in transgenic plants. Concerns of insect resistance development and the desire to achieve the optimum toxicity and control of additional insect species resulted in efforts to modify existing Bt-derived proteins so as to increase their toxicity or alter their mode of action.

Most studies on the mode of action of *Bacillus thuringiensis* toxins have focused on lepidopteran-specific Cry1 insecticidal crystal proteins ("ICPs"). The following picture has emerged from these studies (Gill et al., 1992, Annu. Rev. Entomol. 37, 615-36; Knowles, 1993, BioEssays, 15, 469-476). Following ingestion of the crystals by a susceptible insect, they are dissolved in the alkaline reducing environment of the insect midgut lumen. The liberated proteins, the protoxins, are then proteolytically processed by insect midgut proteases to a protease-resistant fragment. This active fragment, the toxin, then passes through the peritrophic membrane and binds to specific receptors located on the brush border membrane of gut epithelial cells. Subsequent to binding, the toxin or part thereof inserts in the membrane resulting in the formation of pores. These pores lead to colloid osmotic swelling and ultimately lysis of the midgut cells, causing death of the insect.

Binding studies have demonstrated that receptor binding is a crucial step in the mode of action of ICPs (Hofmann et al., 1988, 173, 85-91; Hofmann et al., 1988, Proc. Natl. Acad. Sci. USA, 85, 7844-7848; Van Rie et al., 1990, Appl. Environm. Microbiol. 56, 1378-85).

The three dimensional structure of two ICPs, Cry3A and the Cry1Aa toxic fragment, has been solved (Li et al., 1991, Nature 353, 815-21; Grochulski et al., 1994, Journal of Molecular Biology 254, 1-18). The Cry proteins have been found to have three structural domains: the N-terminal domain I consists of 7 alpha helices, domain II contains three beta-sheets and the C-terminal domain III is a beta-sandwich. Based on this structure, a hypothesis has been formulated regarding the structure-function relationships of ICPs. The bundle of long, hydrophobic and amphipathic helices (domain I) is equipped for pore formation in the insect membrane, and regions of the three-sheet domain (domain II) are probably responsible for receptor binding (Li et al, 1991, *supra*). The function of domain III is less clear. When different ICP amino acid sequences are aligned, five conserved sequence blocks are evident (Höfte & Whiteley, 1989, Microbiol. Revs. 53, 242-255). These conserved blocks are all located in the interior of a structural domain or at the interface between domains. The high degree of conservation of these internal residues implies that homologous proteins would adopt a similar fold (Li et al., 1991, *supra*).

Data from Ahmad et al. (1991, FEMS Microbiol. Lett. 68, 97-104); Wu et al. (1992, J. Biol. Chem. 267, 2311-2317) and Gazit et al. (1993, Biochemistry 32, 3429-3436) provide evidence for the function of domain I of ICPs as a pore formation unit.

Deletions and alanine substitutions in the Cry1Aa protoxin at a position predicted to be at or near the second loop of domain II significantly altered toxicity and receptor binding ability (Lu et al., 1993, XXVIth Annual meeting of the Society for Invertebrate Pathology, Asheville, USA, Conference book, page 31, Abstract 17). Smith and Ellar (1992, XXVth Annual meeting of the Society for Invertebrate Pathology, Heidelberg, Germany, Conference book, page 111, abstract 68) observed dramatic effects on toxicity towards *in vitro* insect cell cultures with mutant Cry1C proteins, differing in the amino acid sequence of the predicted loop regions.

They formulated the hypothesis that it should be possible to map the putative receptor binding domain of this toxin and eventually generate toxins with increased potency. In some cases however, a contribution to specificity and binding from domain III of the Cry toxin could not be excluded (Schnepf et al., 1990, *supra*; Ge et al., 1991, J. Biol. Chem. 266, 17954-17958). Furthermore, a recent study using hybrid ICPs, constructed by exchanging gene fragments between *cry1C* and *cry1E*, has indicated that domain II of Cry1C is not sufficient to confer the high activity of this protein towards *Spodoptera exigua* and *Mamestra brassicae* (Schipper et al., 1993, Seventh International Conference on *Bacillus*, Institut Pasteur, July 18-23, Abstracts of lectures, p. L69). Site-directed mutagenesis experiments on Cry1Ac indicated that certain amino acids in domain I are important for receptor binding (Wu et al., 1992, *supra*). Rajamohan et al. (1996, J. Biol. Chem. 271, 2390-2396) explored the role of loop 2 residues in domain II of the Cry1Ab protein in reversible and irreversible binding to *Manduca sexta* and *Heliothis virescens*.

Also, changes outside the 60 kD toxin region of the Bt protoxin were found to influence toxicity. It was suggested that this may be related to the activation processes by the gut juice (Nakamura et al., 1990, Agric. Biol. Chem. 54, 715-24).

Visser et al. (1993, In *"Bacillus thuringiensis, an Environmental Biopesticide : Theory and Practice"*, pp.71-88, eds.: Entwistle, P.F., Cory, J.S., Bailey, M.J., and Higgs, S., John Wiley & Sons, NY) reviewed the domain-function studies with Bt ICPs and concluded that in general, the function of essential stretches of the toxic fragment of Bt ICPs is unknown. From studies of mutant proteins, it was found that several amino acid residues from different regions of the toxic fragment, either conserved or variable, were shown to affect toxic activity.

Lambert et al. (1996, Appl. & Environm. Microbiol. 62, p. 80-86) and PCT patent publication WO 94/05771 describe a new Bt protein which is currently named cry9Ca1 (abbreviated as Cry9C) (Peferoen et al., 1997, in *Advances in Insect Control: The role of transgenic plants*, pp. 21-48, Taylor & Francis Ltd., London). This protein was found to have a broad insect target range within the group of lepidopteran pest insects making it interesting for insect control applications in agriculture.

De Roeck et al. (1995, the 28th annual meeting of the Society for Invertebrate Pathology, Cornell University, Ithaca, New York, p. 52) suggests to determine the likely position of the binding epitope of the CryIH protein by making Alanine mutants so as to allow the determination of the contribution of amino acid positions in binding of the CryIH protein to different insects. The CryIH protein is currently named Cry9C in the new nomenclature (Crickmore et al., 1995, 28th annual meeting of the Society for Invertebrate Pathology, Cornell University, Ithaca, New York, p.14.). De Roeck et al. (1997, the 6th International Conference on Perspectives in Protein Engineering, John Innes Centre, Norwich, UK, June 28-July 1, p. 34) determined the likely position of residues in the loops at the apex of the molecule in domain II of the Cry9C protein.

SUMMARY OF THE INVENTION

This invention provides a modified Cry9C protein with an improved toxicity to an insect species, comprising the amino acid sequence of SEQ ID No. 2 or an insecticidally-effective fragment thereof, wherein at least one amino acid in the following regions in SEQ ID No. 2 is replaced by another amino acid: 313-334, 358-369, 418-425, 480-492.

This invention further provides improved Cry9C proteins comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein at least one of the amino acids at the following positions in SEQ ID No. 2 have been replaced by another amino acid: 313, 316, 317, 318, 319, 321, 323, 325, 329, 330, 368, 369, 418, 420, 421, 422, 480, 481, 483, 484, 485, 487, 488, 490 and 491. Preferred improved Cry9C proteins comprise the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein at least one of the amino acids at the following positions are replaced by other amino acids: 316, 317, 319, 321, 329, 330, 369, 422, and 488.

This invention also provides a modified Cry9C protein with improved toxicity to *Ostrinia nubilalis*, comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein at least the amino acids at position 488 or at least at positions 364 and 488 are replaced by other amino acids, preferably by alanine.

This invention also provides modified Cry9C proteins with improved toxicity to *Heliothis virescens,* comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein the amino acid at position 321 or position 329, is replaced by another amino acid, preferably by alanine.

This invention further provides modified Cry9C proteins with improved toxicity to *Diatraea grandiosella,* comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein the amino acid at any or all of positions 316, 317, 319, 321, 330, 369, or 422 is replaced by another amino acid, preferably by alanine.

Further in accordance are provided DNA sequences encoding the modified Cry9C proteins, and particularly chimeric genes designed for expression in plants comprising these DNA sequences.

In another preferred embodiment of this invention, a plant transformed with a DNA sequence encoding a modified Cry9C protein is provided, so that the plant acquires increased resistance to insects, particularly a corn plant transformed with a modified Cry9C protein yielding increased toxicity towards *Heliothis virescens*, *Ostrinia nubilalis*, or *Diatraea grandiosella* insects.

Other objects and advantages of this invention will become evident from the following description.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

In this invention, certain amino acid residues important for toxicity of the Cry9C protein have been identified. These amino acid residues can be replaced by other amino acids to increase the toxicity to a specific insect species.

The "Cry9C protein", as used herein, refers to an insecticidal protein characterized by the amino acid sequence of SEQ ID No. 2 or any equivalents thereof such as the insecticidally effective truncated proteins or the fusion proteins of the Cry9C protein described in PCT patent publications WO 94/05771 and WO 94/24264. Particularly preferred Cry9C proteins, in accordance with this invention, are proteins containing at least the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658. Throughout the description

and the claims, the new nomenclature for Bt crystal proteins as suggested by Crickmore et al. (1995, 28th annual meeting of the Society for Invertebrate Pathology, Cornell University, Ithaca, New York, p. 14) and reported in Peferoen et al. (1997, in *Advances in Insect Control: The role of transgenic plants*, pp. 21-48, Taylor & Francis Ltd., London) has been used.

"Cry9C protein variants", for a particular insect species, are insecticidal proteins that differ from but are indirectly or directly derived from the Cry9C protein. Indeed, several variants of a Bt protein in which some amino acids are changed into others without significantly changing activity and/or specificity to a particular insect species can be found in nature (Höfte & Whiteley, 1989, *supra*) or can be made by recombinant DNA techniques. Variants of a Cry9C protein, as used herein, also include proteins containing the specificity- or toxicity-determining domain or region of the Cry9C protein, e.g., in a hybrid with another protein, such as another Bt ICP, a membrane-permeating protein domain, a cytotoxin or an antibody fragment, provided that the Cry9C specificity- or toxicity-determining domain or region contributes to the toxicity or specificity of the hybrid protein. Particularly preferred Cry9C protein variants are those proteins comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein the arginine at position 164 has been replaced by another amino acid, preferably alanine or lysine. These variants with a replacement of the arginine at position 164 in the sequence of SEQ ID No. 2 show a significantly lower susceptibility to breakdown upon protease treatment, and are named herein the "protease-resistant Cry9C variants". Like here for the protease-resistant variants, whenever reference to a particular region or position in SEQ ID No. 2 is made, this does not necessarily imply that the protein referred to is the full-length protein of SEQ ID No. 2; this statement merely refers to the position corresponding to the particular position in the reference Cry9C protein in SEQ ID No. 2. Indeed, improved Cry9C proteins of the invention can be truncated so that the actual position of an amino acid in that protein will differ but nevertheless reference will be made throughout this invention to the positions in the full-length reference protein, shown in SEQ ID No. 2.

Following the teachings of this invention, Cry9C proteins or variants thereof can be modified to have an increased toxicity for an insect species. "Modified Cry9C

protein", as used herein, refers to a Cry9C protein or its protease-resistant variant wherein amino acids have been modified to analyse the contribution of amino acid positions in toxicity, particularly a Cry9C protein or its protease-resistant variant wherein amino acids have been modified in the regions at the following positions in SEQ ID No. 2: 313-334, 358-369, 418-425, 480-492. "Improved Cry9C protein", in accordance with this invention, refers to a Cry9C protein or its protease-resistant variant wherein at least one amino acid has been replaced, so that the toxicity of this improved protein towards an insect species is significantly increased. In a particularly preferred improved Cry9C protein or its protease-resistant variant, the at least one amino acid change is located in domain II of the Cry9C protein, particularly in the regions of the Cry9C protein characterized by the following positions in SEQ ID No. 2: 313-334, 358-369, 418-425, 480-492. A modified Cry9C protein, differing in one amino acid from the native protein or its protease-resistant variant and being significantly less toxic towards the target insect, allows the direct identification of this amino acid position as involved in toxicity (provided no gross structural changes are introduced), and thus has considerable value in improving toxicity. In accordance with this invention, the identification of these amino acid positions involved in toxicity allows the construction of modified proteins having increased toxicity to the target insect by amino acid randomization at these positions. Preferred modified Cry9C proteins in accordance with this invention are the modified Cry9C proteins having altered toxicity to *Ostrinia nubilalis*, *Heliothis virescens* or *Diatraea grandiosella* as shown in Table 1, as well as combinations of those modifications in one modified protein.

An example of an improved Cry9C protein in accordance with this invention is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 or 666 wherein an amino acid in at least one of the following amino acid positions of SEQ ID No. 2 has been replaced by another amino acid: 313, 316, 317, 318, 319, 321, 323, 325, 329, 330, 362, 364, 368, 369, 418, 420, 421, 422, 480, 481, 483, 484, 485, 487, 488, 490 and 491; or an amino acid position located in the immediate vicinity of any one of these positions in the three-dimensional structure of the protein, preferably those amino acids whose C-alpha atom is at a maximum distance of about 7 Angstrom from the C-alpha atom

of the amino acid listed above. A preferred improved Cry9C protein in accordance with this invention is the protein of SEQ ID No. 2 with at least one of the following amino acid changes: P316A, A317V, V319A, L321A, P329A, Y330A, S364A, Y369A, I422A, and I488A. "V319A" or "Cry9C(V319A)", as used herein, means a change of the valine amino acid at position 319 in SEQ ID No. 2 to an alanine amino acid. Preferred improved Cry9C proteins also include Cry9C proteins having also the arginine amino acid at position 164 in SEQ ID No. 2 altered into another amino acid, particularly alanine or lysine, to enhance stability upon protease, particularly trypsin, cleavage.

A preferred Cry9C protein for the control of *Ostrinia nubilalis* insects in accordance with this invention is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein an amino acid in at least one of the following amino acid positions in SEQ ID No. 2 has been replaced by another amino acid: 325, 364, 418, 421, 485, and 488. A particularly preferred improved Cry9C protein for the control of *Ostrinia nubilalis* insects is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein the amino acids in at least position 364 or at least in positions 364 and 488 of SEQ ID No. 2 are replaced by another amino acid, particularly alanine.

A preferred Cry9C protein for the control of *Heliothis virescens* insects in accordance with this invention is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein an amino acid in at least one of the following amino acid positions in SEQ ID No. 2 has been replaced by another amino acid: 313, 316, 317, 318, 319, 321, 323, 325, 329, 330, 368, 369, 418, 420, 421, 422, 480, 481, 483, 484, 485, 487, 488, 490 and 492, particularly at least one of the following amino acid positions: 321, 325, 329, 418, 420, and 480. A particularly preferred improved Cry9C protein for the control of *Heliothis virescens* insects is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein the amino acids in at least one of the amino acid positions 321 and 329 of SEQ ID No. 2 are replaced by another amino acid, particularly alanine.

A preferred Cry9C protein for the control of *Diatraea grandiosella* insects in accordance with this invention is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein an amino acid in at least one of the following amino acid positions in SEQ ID No. 2 has been replaced by another amino acid: 316, 317, 319, 321, 325, 330, 369, 421, 422, 480, 483, 484, 485, 487, 488, 490, and 491; particularly at least one of the following amino acid positions: 480, 484, 485, 487, and 490. A particularly preferred improved Cry9C protein for the control of *Diatraea grandiosella* insects is a protein comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658 wherein the amino acids in at least one of the amino acid positions 316, 317, 319, 321, 330, 369 and 422 of SEQ ID No. 2 are replaced by another amino acid, particularly alanine or valine (for 317).

By using DNA sequences encoding improved Cry9C proteins in accordance with this invention, improved toxicity to a selected insect species can be obtained upon expression of such DNA in a transgenic plant.

A "*cry9C* gene", as used herein, is a DNA sequence comprising a DNA encoding a Cry9C protein (a coding region), and includes necessary regulatory sequences so that a Cry9C protein can be expressed in a cell, preferably a plant or bacterial cell. A *cry9C* gene does not necessarily need to be expressed everywhere at all times, expression can be periodic (e.g. at certain stages of development in a plant) and/or can be spatially restricted (e.g. in certain cells or tissues in a plant), mainly depending on the activity of regulatory elements provided in the chimeric gene or in the site of insertion in the plant genome. A *cry9C* gene can be naturally-occurring or can be a hybrid or synthetic DNA and the regulatory elements can be from prokaryotic or eucaryotic origin.

The "modified *cry9C* gene", as used herein, is a DNA sequence comprising a DNA encoding a modified Cry9C protein (a modified coding region), and includes necessary regulatory sequences so that a Cry9C protein can be expressed in a cell, preferably a plant or bacterial cell. An example of a modified *cry9C* coding region is the *cry9C* coding region of SEQ ID No. 3 wherein the valine codon at nucleotide positions 844-846 of SEQ ID No. 3 has been replaced by an alanine codon.

"Substantial sequence homology" to a DNA sequence, as used herein, refers to DNA sequences differing in some, most or all of their codons from another DNA sequence but encoding the same or substantially the same protein. Indeed, because of the degeneracy of the genetic code, the codon usage of a particular DNA coding region can be substantially modified, e.g., so as to more closely resemble the codon usage of the genes in the host cell, without changing the encoded protein. Changing the codon usage of a DNA coding region to that of the host cell has been described to be desired for gene expression in foreign hosts (e.g. Bennetzen & Hall, 1982, J. Biol. Chem. *257*, 3026-3031.; Itakura, 1977, Science *198*, 1056-1063). Codon usage tables are available in the literature (Wada et al., 1990, Nucl. Acids Res. *18*, 2367-1411; Murray et al., 1989, Nucl. Acids Res. *17*(2), 477-498) and in the major DNA sequence databanks (e.g. at EMBL in Heidelberg, Germany). Accordingly, recombinant or synthetic DNA sequences can be constructed so that the same or substantially the same proteins with substantially the same insecticidal activity are produced (Koziel et al., 1993, Bio/technology 11, 194-200; Perlak et al., 1993, Plant Mol. Biol. 22, 313-321). A modified *cry9C* gene has all appropriate control regions so that the modified Cry9C protein can be expressed in a host cell, e.g. for expression in plants, a plant-expressible promoter and a 3' termination and polyadenylation region active in plants.

A "chimeric improved *cry9C* gene", as used herein, refers to a chimeric gene comprising a DNA sequence encoding the improved Cry9C protein inserted in between controlling elements of different origin, e.g. a DNA sequence encoding the improved Cry9C protein under the control of a promoter transcribing the DNA in the plant cell, and fused to 3' transcription termination sequences active in plant cells.

Protection of a plant, preferably a corn or cotton plant, against an insect species which is known to feed on said plant is preferably accomplished by expressing an improved Cry9C protein in the cells of the plant. This is preferably accomplished by expressing a chimeric improved *cry9C* gene encoding such an improved Cry9C protein in the cells of a plant, preferably a corn or cotton plant. An improved Cry9C protein of this invention preferably only has a small number, particularly less than 20, more particularly less than 15, preferably less than 10 amino acids replaced by other amino acids as compared to the Cry9C protein,

preferably as compared to the region from between amino acid positions 1 and 45 to amino acid position 658 of the Cry9C protein of SEQ ID No. 2. A significant increase in toxicity can already be obtained by replacing only 1 amino acid, but it is preferred that more than one amino acid is changed to improve toxicity.

The following steps are followed to construct the new modified Cry9C proteins: amino acids in domain II of the Cry9C protein from amino acid positions 313-334, 358-369, 418-425, and 480-492 were chosen for modification, using alanine-scanning mutagenesis (Cunningham & Wells, 1989, Science 244, 1081-85). In case the original position is alanine, a substitution by valine is done. These regions occur at positions corresponding to the solvent-exposed positions in the loop between beta-strands 1 and 2 (comprising alpha-helix 8) and in loop 1 (located between beta strands 2 and 3), in loop 2 (located between beta-strands 6 and 7), and in loop 3 (located between beta-strands 10 and 11) in the three-dimensional model of the Cry3A protein (Li et al., 1991, *supra*). To discount any observed lower toxicity of a modified Cry9C protein which is due to misfolding or structural distortion, the structural stability of mutant ICPs can be analysed by a variety of methods including toxicity to another target insect, crystal formation, solubilization, monoclonal antibody binding analysis, protease resistance, fluorometric monitoring of unfolding and circular dichroism spectrum analysis. In the case of structural distortion, it is impossible to determine the functional role of this position by alanine replacement. However, a more conservative amino acid substitution may yield a correctly folded mutant protein which allows to determine the functional role of this position.

The amino acid positions, identified above, which yield modified proteins with significantly decreased toxicity ("down-mutants") are randomized. This means that a set of 20 different mutants, representing each type of amino acid, is generated for each position of interest (the original amino acid and the alanine substitution function as a control). This method is further referred to as "amino acid randomization". Such mutants may be generated by a variety of methods, e.g. following the PCR overlap extension method (Ho et al., 1989, Gene 77, 51-59). These mutant proteins are then tested in toxicity assays on the target insect. Mutants at each position which are more toxic, e.g., yield higher mortality than the wild type protein, are selected. Such mutants with improved toxicity are termed "up-mutants". Alternatively, it is also

possible to select potential up-mutants on the basis of increased reversible binding which can be measured following the procedures of Van Rie et al. (1990, Appl. Environm. Microbiol. 56, 1378-1385) or Liang et al. (1995, J. Biol. Chem. 270, 24719-24724), which is incorporated herein by reference.

All or some of the "up-mutant" amino acids, identified in step 2, are combined in a single modified protein. According to additivity principles, mutations in non-interacting parts of a protein should combine to give simple additive changes in the free energy of binding (Lowman and Wells, 1993, J. Mol. Biol., 234, 564-578). Increases in toxicity are thus accumulated by combining several single mutants into one multiple mutant. Finally a modified protein with improved toxicity is designed, which comprises some or all, preferably all, of the up-mutant amino acids previously identified.

In accordance with this invention, amino acids of domain II of a Cry9C protein, located at the protruding regions of domain II are chosen for modification. By "protruding regions of domain II", as used herein, are meant the solvent-exposed regions organized in loops, alpha helices or beta-strands which are protruding from domain II and are located at or towards the apex of the molecule.

This invention is particularly suited for improving the toxicity to an insect species for which the Cry9C protein has a rather weak toxicity. The toxicity of this improved Cry9C protein can be increased by combining amino acid mutations in the protein, each yielding an increased toxicity when compared to the amino acid present in the native Cry9C protein. Insect species for which improved Cry9C proteins can be made also include *Spodoptera frugiperda*, *Heliothis zea*, *Heliothis armigera*, and *Agrotis ipsilon*. Also, this invention is suited to increase toxicity of a Cry9C protein or its protease-resistant variant to one insect species and to decrease toxicity of the same protein to another insect species by making the proper amino acid substitutions in the protein. This may be advantageous, e.g., to limit the likelihood of insect resistance occurrence to the protein in a particular insect species.

An insecticidally effective part of the modified *cry9C* gene of this invention encoding an insecticidally effective portion of the modified Cry9C protein, can be made in a conventional manner. An "insecticidally effective part" of the modified

cry9C gene refers to a gene comprising a DNA coding region encoding a polypeptide with fewer amino acids than the full length modified Cry9C protein but that still retains toxicity to insects. A preferred insecticidally effective part of the Cry9C protein is the part from amino acid position 1 or 44 to amino acid position 658 in SEQ ID No. 2.

In order to express all or an insecticidally effective part of the improved *cry9C* gene in *E. coli*, in Bt strains and in plants, suitable restriction sites can be introduced, flanking each gene or gene part. This can be done by site-directed mutagenesis, using well-known procedures (Stanssens et al., 1989, Nucl. Acids Res. *12*, 4441-4454; White et al., 1989, Trends in Genet. 5, 185-189).

In order to improve expression in foreign host cells such as plant cells, it may be preferred to alter the improved *cry9C* coding region or its insecticidally effective part to form an equivalent, artificial improved *cry9C* coding region. Expression is improved by selectively inactivating certain cryptic regulatory or processing elements present in the native sequence as described in PCT publications WO 91/16432 and WO 93/09218. This can be done by site-directed mutagenesis or site-directed intron-insertion (WO 93/09218), or by introducing overall changes to the codon usage, e.g., adapting the codon usage to that most preferred by the host organism (publication of European patent application number ("EP") 0 385 962, EP 0 359 472, publication of PCT patent application WO 93/07278, Murray et al., 1989, *supra*) without significantly changing, preferably without changing, the encoded amino acid sequence. Small modifications to a DNA sequence such as described above can be routinely made by PCR-mediated mutagenesis (Ho et al., 1989, *supra*; White et al., 1989, *supra*). For major changes to the DNA sequence, DNA synthesis methods are available in the art (e.g. Davies et al., 1991, Society for Applied Bacteriology, Technical Series 28, pp. 351-359). For obtaining enhanced expression in monocot plants such as corn, a monocot intron can be added to the chimeric improved *cry9C* gene (Callis et al., 1987, Genes & Development 1, 1183-1200; PCT publication WO 93/07278). Another preferred embodiment of this invention is the expression of the improved Cry9C proteins by the method described in PCT patent publication WO 97/49814, which is incorporated herein by reference.

The chimeric improved *cry9C* gene can be stably inserted in a conventional manner into the nuclear genome of a single plant cell, and the so-transformed plant

cell can be used in a conventional manner to produce a transformed plant that is insect-resistant. Particularly preferred plants in accordance with this invention are corn plants. Corn cells can be stably transformed (e.g. by electroporation) using wounded or enzyme-degraded intact tissues capable of forming compact embryogenic callus (such as corn immature embryos), or the embryogenic callus (such as type I callus in corn) obtained thereof, as described in PCT patent publication WO 92/09696 or US Patent 5,641,664. Other methods for transformation of corn include the methods by Fromm et al. (1990, Bio/Technology *8*, 833-839), Gordon-Kamm et al. (1990, The Plant Cell 2, 603-618) and Ishida et al. (1996, Nature Biotechnology 14, 745-750).

Alternatively, a disarmed Ti plasmid, containing the insecticidally effective chimeric improved *cry9C* gene, in *Agrobacterium tumefaciens* can be used to transform the plant cell, preferably the corn or cotton cell, and thereafter, a transformed plant can be regenerated from the transformed plant cell using the procedures described, for example, in EP 0116718, EP 0270822, PCT publication WO 84/02913 and EP 0242246 (which are also incorporated herein by reference), and in Gould et al. (1991, Plant Physiol. *95*, 426-434) or Ishida et al. (1996, *supra*), particularly the method described in PCT publication WO 94/00977. Preferred Ti-plasmid vectors each contain the insecticidally effective chimeric improved *cry9C* gene between the border sequences, or at least located to the left of the right border sequence, of the T-DNA of the Ti-plasmid. Of course, other types of vectors can be used to transform the plant cell, using procedures such as direct gene transfer (as described, for example in EP 0233247), pollen mediated transformation (as described, for example in EP 0270356, PCT publication WO 85/01856, and US Patent 4,684,611), plant RNA virus-mediated transformation (as described, for example in EP 0067553 and US Patent 4,407,956), and liposome-mediated transformation (as described, for example in US Patent 4,536,475).

A resulting transformed plant, such as a transformed corn or cotton plant, can be used in a conventional plant breeding scheme to produce more transformed plants with the same characteristics or to introduce the improved *cry9C* gene, or an insecticidally effective part thereof in other varieties of the same or related plant species. Seeds, which are obtained from the transformed plants, contain the

chimeric improved *cry9C* gene or its insecticidally effective part as a stable genomic insert. Cells of the transformed plant can be cultured in a conventional manner to produce the improved Cry9C protein or insecticidally effective portions thereof, which can be recovered for use in conventional insecticide compositions against insects, particularly lepidopteran insects (U.S. Patent 5,254,799). Preferred plants in accordance with this invention, besides corn and cotton, include rice, plants of the genus Brassica such as oilseed rape, cauliflower and broccoli, and also soybean, tomato, tobacco, potato, eggplant, beet, oat, pepper, gladiolus, dahlia, chrysanthemum, sorghum, and garden peas.

The improved *cry9C* coding region or its insecticidally effective part is inserted in a plant cell genome so that the inserted coding region is downstream (i.e., 3') of, and under the control of, a promoter which can direct the expression of the gene part in the plant cell. This is preferably accomplished by inserting the chimeric improved *cry9C* gene or its insecticidally effective part in the plant cell genome. Preferred promoters include: the strong constitutive 35S promoters (the "35S promoters") of the cauliflower mosaic virus of isolates CM 1841 (Gardner et al., 1981, Nucleic Acids Research *9*, 2871-2887), CabbB-S (Franck et al., 1980, Cell *21*, 285-294) and CabbB-JI (Hull and Howell, 1987, Virology *86*, 482-493); the ubiquitin promoter (EP 0342926), and the TR1' promoter and the TR2' promoter which drive the expression of the 1' and 2' genes, respectively, of the T-DNA (Velten et al., 1984, EMBO J. *3*, 2723-2730). Alternatively, a promoter can be utilized which is not constitutive but rather is specific for one or more tissues or organs of the plant, preferably leaf and stem tissue, whereby the inserted chimeric improved *cry9C* gene or its insecticidally effective part is expressed only in cells of the specific tissue(s) or organ(s). Another alternative is to use a promoter whose expression is inducible (e.g., by insect feeding or by chemical factors). Known wound-induced promoters inducing systemic expression of their gene product throughout the plant are also of particular interest.

The improved *cry9C* coding region, or its insecticidally effective part, is inserted in the plant genome so that the inserted coding region is upstream (i.e., 5') of suitable 3' end transcription regulation signals (i.e., transcript termination and polyadenylation signals). Preferred polyadenylation and transcript formation signals include those of the 35S gene (Mogen et al., 1990, The Plant Cell 2, 1261-1272), the

octopine synthase gene (Gielen et al., 1984, EMBO J. *3*, 835-845) and the T-DNA gene 7 (Velten and Schell, 1985, Nucl. Acids Res. *13*, 6981-6998), which act as 3'-untranslated DNA sequences in transformed plant cells.

The chimeric improved *cry9C* gene, or its insecticidally effective gene part, can optionally be inserted in the plant genome as a hybrid gene (EP 0 193 259; Vaeck et al., 1987, Nature 327, 33-37) under the control of the same promoter as the coding region of a selectable marker gene, such as the coding region of the *neo* gene (EP 0 242 236) encoding kanamycin resistance, so that the plant expresses a fusion protein.

Preferably, the improved *cry9C* gene is expressed in a plant in combination with another insect control protein, e.g., another Bt-derived crystal protein or an insecticidal fragment thereof, particularly a Cry1Ab- or Cry1B-type protein, to prevent or delay the occurrence of insect resistance development (EP 0 408 403).

All or part of the improved *cry9C* coding region can also be used to transform bacteria, such as a *B. thuringiensis* which produces other insecticidal toxins (Lereclus et al., 1992, Bio/Technology 10, 418-421; Gelernter & Schwab, 1993, In *Bacillus thuringiensis, An Environmental Biopesticide: theory and Practice*, pp. 89-104, eds. Entwistle, P.F., Cory, J.S., Bailey, M.J. and Higgs, S.; John Wiley & Sons Ltd.). Thereby, a transformed Bt strain is produced which is useful for combating a wide spectrum of insect pests or for combating insects in such a manner that insect resistance development is prevented or delayed (EP 0 408 403). Preferred promoter and 3' termination and polyadenylation sequences for the chimeric improved *cry9C* gene are derived from *Bacillus thuringiensis* genes, such as the native ICP genes.

Alternatively, the improved coding region of the invention can be inserted and expressed in endophytic and/or root-colonizing bacteria, such as bacteria of the genus *Pseudomonas* or *Clavibacter*, e.g., under the control of a Bt ICP gene promoter and 3' termination sequences. Successful transfer and expression of ICP genes into such bacteria has been described by Stock et al. (1990, Can. J. Microbiol. 36, 879-884), Dimock et al. (1989, In *Biotechnology, Biopesticides and Novel Plant Pest Resistance Management*, eds. Roberts, D.W. & Granados, R.R., pp.88-92, Boyce Thompson Institute for Plant Research, Ithaca, New York), and Waalwijk et al. (1991, FEMS Microbiol. Lett. 77, 257-264). Transformation of bacteria with all or part

of the improved *cry9C* coding region of the invention, incorporated in a suitable cloning vehicle, can be carried out in a conventional manner, preferably using conventional electroporation techniques as described in Mahillon et al. (1989, FEMS Microbiol. Letters *60*, 205-210), in PCT patent publication WO 90/06999, Chassy et al. (1988, Trends Biotechnol. *6*, 303-309) or other methods, e.g., as described by Lereclus et al. (1992, Bio/Technology *10*, 418).

The improved Cry9C-producing strain can also be transformed with all or an insecticidally effective part of one or more DNA sequences encoding a Bt protein or an insecticidally effective part thereof, such as: a DNA encoding the Bt2 or Cry1Ab protein (US patent 5,254,799; EP 0 193 259) or the Bt109P or Cry3C protein (PCT publication WO 91/16433), or another DNA coding for an anti-lepidoptera or an anti-Coleoptera protein. Thereby, a transformed Bt strain can be produced which is useful for combating an even greater variety of insect pests (e.g., Coleoptera and/or additional lepidoptera) or for preventing or delaying the development of insect resistance.

For the purpose of combating insects by contacting them with the improved Cry9C protein, e.g. in the form of transformed plants or insecticidal formulations, any DNA sequence encoding any of the above described improved Cry9C proteins, can be used.

The following Examples are offered by way of illustration and not by way of limitation. The sequence listing referred to in the description and the Examples is as follows:

SEQUENCE LISTING

SEQ ID No. 1: Nucleotide sequence of the *Bacillus thuringiensis cry9C* gene, showing the coding region and flanking 5' and 3' regions.

SEQ ID No. 2: Amino acid sequence of the full length *Bacillus thuringiensis* Cry9C protein.

SEQ ID No. 3: Nucleotide sequence of a codon-optimized DNA sequence encoding a truncated Cry9C protein wherein the arginine at

amino acid position 123 (corresponding to amino acid position 164 in the protein of SEQ ID No. 2) has been replaced by lysine.

SEQ ID No. 4: Amino acid sequence of the modified Cry9C protein encoded by the DNA of SEQ ID No. 3.

Unless otherwise stated in the Examples, all general materials and methods, including procedures for making and manipulating recombinant DNA are carried out by the standardized procedures as described in volumes 1 and 2 of Ausubel et al., *Current Protocols in Molecular Biology*, Current Protocols, USA (1994), in Plant Molecular Biology Labfax (1993, by R.D.D. Croy, jointly published by BIOS Scientific publications Ltd. UK and Blackwell Scientific Publications, UK) and Sambrook et al., *Molecular Cloning - A Laboratory Manual, Second Ed.*, Cold Spring Harbor Laboratory Press, NY (1989).

EXAMPLES:

1. CONSTRUCTION OF MODIFIED CRY9C PROTEINS

Multiple alignments between Bt crystal protein sequences including the sequences of Cry9C, Cry3A and Cry1Aa allowed identification of the amino acids located in the expected binding site of the Cry9C domain II. Using known alignment programs, 52 amino acid positions were identified for amino acid replacement. The amino acids in the Cry9C protein of SEQ ID No. 2 from amino acid positions 313-334, 358-369, 418-425, 480-492 have been identified to correspond to the solvent-accessible regions most likely involved in receptor-binding in the Cry3A protein; and these positions in the Cry9C protein were chosen for amino acid modification. Since alanine substitution does not alter the main chain of a protein, and does not impose extreme electrostatic or steric effects and since it eliminates the side chain beyond the beta carbon, each of the amino acids in these identified regions was changed into alanine, one by one, using splice overlap extension PCR (Ho et al., 1989, *supra*) on the protease-resistant form of the native *cry9C* gene wherein the arginine codon at position 164 was replaced by an alanine codon. The codon most preferred in the *cry9C* native gene for alanine, GCA, was used for these modifications. When the original codon encodes alanine, then this is replaced by a valine codon (GTA). The

obtained PCR fragments were ligated in pUC19-derived vectors. If not present, suitable unique restriction sites were created in the *cry9C* DNA. All plasmids containing modified DNA sequences were controlled by sequencing the relevant portions and were found to be correctly constructed. The modified *cry9C* genes were expressed in transformed WK6 cells. Every mutant protein was expressed in these *E. coli* cells at least twice. Mutants causing problems in expression, probably caused by structural changes in these mutants, were discarded. No gross folding aberrations of the mutants identified to be involved in toxicity (and listed in Table 1) are found, e.g., as was evidenced by the similar SDS-PAGE patterns following trypsin cleavage or treatment with midgut juice of the insect larvae of solubilized mutant and Cry9C(R164A) proteins.

2. INSECT TOXICITY OF THE MODIFIED CRY9C PROTEINS

Bio assays on the modified Cry9C proteins obtained in Example 1 were carried out with first instar larvae of the Southwestern corn borer, *Diatraea grandiosella* (family Pyralidae); the European corn borer, *Ostrinia nubilalis* (family Pyralidae); and the tobacco budworm, *Heliothis virescens* (family Noctuidae). A dilution series of each protein was surface-layered on the artificial diet to determine the LC_{50} value. The artificial diet consisted of: agar (20 g), water (1,000 ml), corn flour (96 g, ICN Biochemicals), yeast (30 g), wheat germs (64 g, ICN Biochemicals), wesson salt (7.5 g, ICN Biochemicals), casein (15 g), sorbic acid (2 g), aureomycin (0.3 g), nipagin (1 g), wheat germ oil (4 ml), sucrose (15 g), cholesterol (1 g), ascorbic acid (3.5 g), Vanderzand modified vitamin mix (12 g, ICN Biochemicals). Larvae were placed on the diet in multi-well plates, 1 larva per well (2 for *Ostrinia nubilalis*). For each dilution, 24 larvae were tested, and dead and living larvae were counted after 5 days. Prior to application, the mutant proteins were digested with trypsin to release the toxin fragments. For each mutant protein, the assays are repeated at least 5 times, using two different protein preparations. As control protein, the trypsin-digested Cry9C(R164A) protein was used. The Cry9C(R164A) protein has the amino acid sequence of SEQ ID No. 2 wherein the arginine at position 164 was replaced by alanine. This protein was found to be more stable than the wild-type Cry9C toxin while retaining its toxicity to the test insects (see, e.g., PCT patent

publication WO 94/24264). The LC_{50} values were calculated with the POLO-program, which is based on the probit analysis (POLO-PC, LeOra Software, 1119 Shattuck Ave., Berkeley California 94707). The results of these assays for those protein mutants which gave an LC_{50} value that is significantly different from that of the control protein in repeated bio assays are summarized in Table 1. It is clear that different positions in the Cry9C protein when substituted to alanine cause increased toxicity in each of the tested insects.

Binding assays on isolated brush border membrane vesicles of *Heliothis virescens* and *Ostrinia nubilalis* performed as described in Van Rie et al. (1990, Appl. Environm. Microbiol. 56, 1378-1385) showed that for all, with the exception of two, of the modified Cry9C proteins with altered toxicity, receptor binding is also altered (e.g., an observed shift in K_D value), thus confirming that for most amino acid residues altered toxicity is due to altered receptor binding. Hence, these residues are proper candidates for improvement of toxicity by amino acid randomization at or near the identified critical position.

3. COMPETITION BINDING EXPERIMENTS

The Cry9C(R164A) protein was tested in competition binding assays using the ECL protein biotinylation system (Amersham Life Sciences, Amersham International plc., UK) as described by Lambert et al. (1996, *supra*) to determine if competition occurred with other Bt toxins in selected insects. For the assays, 3ng biotinylated Cry9C(R164A) protein was added to 30 µg brush border membrane vesicles in PBS buffer (comprising 0,1 % BSA) in the presence of a 300-fold excess of non-biotinylated toxin (homologous competition assays were included in every test as control). Repeated competition tests showed that in both *Ostrinia nubilalis* and *Heliothis virescens* brush border membranes, there was no detectable competition in receptor binding between the (activated) Cry9C(R164A) protein and any one of the following (activated) Bt toxins: the Cry1Aa (Schnepf et al., 1985, J. Biol. Chem. 260, 6264-6272), Cry1Ab (Höfte et al.,1986, Eur. J. Biochem. 161, 271-280), Cry1Ac (Adang et al., 1985, Gene 36, 289-300), Cry1B (Brizzard & Whiteley, 1988, Nucl. Acids Res. 16, 4168-4169) and Cry1C (Honée et al., 1988, Nucl. Acids Res. 16, 6240) toxins. Thus, in these insects the Cry9C(R164A) protein binds to a different

receptor than these other Bt toxins. In *Diatreae grandiosella* competition assays, it was found that the Cry9C(R164A) does compete for a receptor site with the Cry1B and Cry1C Bt toxins, but does not compete with any one of the Cry1Aa, Cry1Ab, and Cry1Ac toxins.

The same results are found for all three insects when testing the Cry9C protein with the amino acid sequence of SEQ ID No. 2 from amino acids 1-658.

Thus, in all these three insects, combination of the Cry9C and a selected non-competitively binding Bt toxin with good toxicity to the target insect can be used simultaneously in order to prevent or delay insect resistance development. In transgenic corn plants, a particularly interesting combination would be the Cry9C (or its protease-resistant variant) and a Cry1B and/or any of the Cry1A-type toxins for *Ostrinia nubilalis* control and the Cry9C (or its protease-resistant variant) and any one of the Cry1A-type toxins, preferably a Cry1Ab-type toxin, for *D. grandiosella* control. For *Heliothis virescens* control, the Cry9C (or its protease-resistant variant) and any of the Cry1A-type toxins are preferred toxins to be co-expressed.

4. CONSTRUCTION OF IMPROVED CRY9C PROTEINS

The modified position in every mutant protein of Example 2 giving rise to a significantly decreased or increased toxicity to an insect species is altered to all other amino acids and the toxicity is re-evaluated. The amino acids yielding the highest toxicity at a particular position are combined to form an improved Cry9C protein. Also the alanine mutants yielding an increase in toxicity (up-mutant amino acid positions) are included in such combinations to form improved Cry9C proteins for the selected insect species. Table 1 indeed shows already two up-mutant proteins for every insect tested. Analysis of all these improved Cry9C proteins in the bio assay shows that combinations of up-mutant amino acid positions can substantially increase toxicity of the Cry9C protein towards selected insect species.

5. GENE CONSTRUCTION AND PLANT TRANSFORMATION

A modified DNA sequence encoding a truncated Cry9C(R164K) protein for expression in corn and cotton plants is shown in SEQ ID No. 3. This DNA sequence has an optimized codon usage for plants and encodes an N- and C-terminally

truncated Cry9C protein wherein an arginine amino acid has been replaced by a lysine (at position 123 in SEQ ID No. 3). Based on this DNA sequence, DNA sequences are made encoding the above improved Cry9C proteins and comprising amino acids 1 to 666 of the Cry9C(R164K) protein. Preferred codons to encode the amino acid replacements in the improved Cry9C proteins are those which are most preferred by the plant host (see, e.g., Murray, 1989, *supra*). A chimeric improved *cry9C* gene comprising the 35S promoter and 35S 3' transcription termination and polyadenylation signal is constructed by routine molecular biology techniques as described in the detailed description.

Corn cells are stably transformed by either *Agrobacterium*-mediated transformation (Ishida et al., 1996, *supra* and U.S. Patent No. 5,591,616) or by electroporation using wounded and enzyme-degraded embryogenic callus, as described in WO 92/09696 or US Patent 5,641,664 (incorporated herein by reference). The resulting transformed cells are selected by means of the incorporated selectable marker gene, grown into plants and tested for susceptibility towards insects. Corn plants expressing a truncated improved Cry9C(R164K) protein wherein the amino acids at positions 364, 488, 319 and 321 have been replaced into alanine show a significantly higher protection from *Ostrinia nubilalis* and *Diatraea grandiosella* damage in comparative tests against corn plants expressing a truncated Cry9C(R164K) protein. A positive correlation is found between the level of expression, as measured by RNA and protein analysis, and the observed insecticidal effect.

Cotton cells are stably transformed by *Agrobacterium*-mediated transformation (Umbeck et al., 1987, Bio/Technology 5, 263-266; US Patent 5,004,863, incorporated herein by reference). The resulting transformed cells are selected by means of the incorporated selectable marker gene, grown into plants and tested for susceptibility towards insects. Cotton plants expressing the truncated improved Cry9C(R164K, L321A, P329A) protein at similar levels than cotton plants expressing the truncated Cry9C(R164K) protein show a significantly higher protection from *Heliothis virescens* damage. A positive correlation is found between the level of expression, as measured by RNA and protein analysis, and the observed insecticidal effect.

The examples and embodiments of this invention described herein are only supplied for illustrative purposes. Many variations and modifications in accordance with the present invention are known to the person skilled in the art and are included in this invention and the scope of the claims. For instance, it is possible to alter, delete or add some nucleotides or amino acids to certain regions of the DNA or protein sequences of the invention without departing from the invention.

All publications (including patent publications) referred to in this application are hereby incorporated by reference, particularly WO 94/05771, WO 94/24264, and Lambert et al. (1996, *supra*).

Table 1: relative toxicity of modified trypsin-digested Cry9C proteins to different insects when compared with the Cry9C(R164A) trypsin-digested protein (mutant 'F313A': the Cry9C(R164A) trypsin-digested protein wherein also the phenylalanine at position 313 is replaced by alanine; 'down(2x)': mutant protein with a significantly lower toxicity (LC50 value about 2 times higher than the control protein), 'up (2x)': mutant with a significantly higher toxicity (LC50 value about two times lower than that of the control protein), '-': no difference in toxicity found):

mutant	*H. virescens*	*O. nubilalis*	*D. grandiosella*
F313A	down (2x)	-	-
P316A	-	-	up (2x)
A317V	-	-	up (2x)
N318A	down (2-3x)	-	-
V319A	-	-	up (3x)
L321A	up (2x)	-	up (2x)
R323A	down (3x)	-	-
W325A	down (4-5x)	down (2x)	down (2-3x)
P329A	up (2x)	-	-
Y330A	-	down (1.5x)	up (2x)
V362A	down (3-4x)	-	-
S364A	-	up (2x)	-
D368A	down (2-3x)	-	-
Y369A	-	-	up (2x)
R418A	down (16x)	down (2x)	-
A420V	down (12x)	-	-
L421A	-	down (2x)	-
I422A	-	-	up (2x)
F480A	down (5x)	-	down (40x)

mutant	*H. virescens*	*O. nubilalis*	*D. grandiosella*
Q481A	down (3x)	-	-
N483A	-	-	down (2x)
Q484A	-	-	down (20x)
A485V	down (3x)	down (2x)	down (20x)
S487A	down (2x)	-	down (20x)
I488A	down (2x)	up (2-3x)	down (5x)
N490A	-	-	down (20x)
A491V	-	-	down (3x)

SEQUENCE LISTING

(1) GENERAL INFORMATION:
(i) APPLICANT:
(A) NAME: PLANT GENETIC SYSTEMS N.V.
(B) STREET: Jozef Plateaustraat 22
(C) CITY: Gent
(E) COUNTRY: Belgium
(F) POSTAL CODE (ZIP): B-9000
(G) TELEPHONE: (32)(9)2358411
(H) TELEFAX: (32)(9)2231923

(ii) TITLE OF INVENTION: Improved Bacillus thuringiensis toxin

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/884,389
(B) FILING DATE: 27-JUN-1997

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4344 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION:668..4141
(D) OTHER INFORMATION:/note= "coding sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCGAGC TCGGTACCTT TTCAGTGTAT CGTTTCCCTT CCATCAGGTT TTCAAATTGA 60

AAAGCCGAAT GATTTGAAAC TTGTTTACGA TGTAAGTCAT TTGTCTATGA CGAAAGATAC 120

GTGTAAAAAA CGTATTGAGA TTGATGAATG TGGACAAGTA GAAATTGACT TACAAGTATT 180

AAAGATTAAG GGTGTCCTTT CTTTTATCGG AAATTTCTCT ATTGAACCTA TTCTGTGTGA 240

AAACATGTAT ACAACGGTTG ATAGAGATCC GTCTATTTCC TTAAGTTTCC AAGATACGGT 300

ATATGTGGAC CATATTTTAA AATATAGCGT CCAACAACTA CCATATTATG TAATTGATGG 360

TGATCATATT CAAGTACGTG ATTTACAAAT CAAACTGATG AAAGAGAATC CGCAATCTGC 420

TCAAGTATCA GGTTTGTTTT GTTTTGTATA TGAGTAAGAA CCGAAGGTTT GTAAAAAAGA 480

AATAGGAATA AATACTATCC ATTTTTTCAA GAAATATTTT TTTATTAGAA AGGAATCTTT 540

CTTACACGGG AAAATCCTAA GATTGAGAGT AAAGATATAT ATATATAAAT ACAATAAAGA 600

GTTTGTCAGG ATTTTTGAAA GATATGATAT GAACATGCAC TAGATTTATA GTATAGGAGG 660

AAAAAGTATG AATCGAAATA ATCAAAATGA ATATGAAATT ATTGATGCCC CCCATTGTGG 720

GTGTCCATCA GATGACGATG TGAGGTATCC TTTGGCAAGT GACCCAAATG CAGCGTTACA 780

AAATATGAAC TATAAAGATT ACTTACAAAT GACAGATGAG GACTACACTG ATTCTTATAT 840

AAATCCTAGT TTATCTATTA GTGGTAGAGA TGCAGTTCAG ACTGCGCTTA CTGTTGTTGG 900

GAGAATACTC GGGGCTTTAG GTGTTCCGTT TTCTGGACAA ATAGTGAGTT TTTATCAATT 960

CCTTTTAAAT ACACTGTGGC CAGTTAATGA TACAGCTATA TGGGAAGCTT TCATGCGACA 1020

GGTGGAGGAA CTTGTCAATC AACAAATAAC AGAATTTGCA AGAAATCAGG CACTTGCAAG 1080

ATTGCAAGGA TTAGGAGACT CTTTTAATGT ATATCAACGT TCCCTTCAAA ATTGGTTGGC 1140

TGATCGAAAT GATACACGAA ATTTAAGTGT TGTTCGTGCT CAATTTATAG CTTTAGACCT 1200

TGATTTTGTT AATGCTATTC CATTGTTTGC AGTAAATGGA CAGCAGGTTC CATTACTGTC 1260

AGTATATGCA CAAGCTGTGA ATTTACATTT GTTATTATTA AAAGATGCAT CTCTTTTTGG 1320

AGAAGGATGG GGATTCACAC AGGGGGAAAT TTCCACATAT TATGACCGTC AATTGGAACT 1380

AACCGCTAAG TACACTAATT ACTGTGAAAC TTGGTATAAT ACAGGTTTAG ATCGTTTAAG 1440

AGGAACAAAT ACTGAAAGTT GGTTAAGATA TCATCAATTC CGTAGAGAAA TGACTTTAGT 1500

GGTATTAGAT GTTGTGGCGC TATTTCCATA TTATGATGTA CGACTTTATC CAACGGGATC 1560

AAACCCACAG CTTACACGTG AGGTATATAC AGATCCGATT GTATTTAATC CACCAGCTAA 1620

TGTTGGACTT TGCCGACGTT GGGGTACTAA TCCCTATAAT ACTTTTTCTG AGCTCGAAAA 1680

TGCCTTCATT CGCCCACCAC ATCTTTTTGA TAGGCTGAAT AGCTTAACAA TCAGCAGTAA 1740

TCGATTTCCA GTTTCATCTA ATTTTATGGA TTATTGGTCA GGACATACGT TACGCCGTAG 1800

TTATCTGAAC GATTCAGCAG TACAAGAAGA TAGTTATGGC CTAATTACAA CCACAAGAGC 1860

AACAATTAAT CCCGGAGTTG ATGGAACAAA CCGCATAGAG TCAACGGCAG TAGATTTTCG 1920

TTCTGCATTG ATAGGTATAT ATGGCGTGAA TAGAGCTTCT TTTGTCCCAG GAGGCTTGTT 1980

TAATGGTACG ACTTCTCCTG CTAATGGAGG ATGTAGAGAT CTCTATGATA CAAATGATGA 2040

ATTACCACCA GATGAAAGTA CCGGAAGTTC AACCCATAGA CTATCTCATG TTACCTTTTT 2100

TAGCTTTCAA ACTAATCAGG CTGGATCTAT AGCTAATGCA GGAAGTGTAC CTACTTATGT 2160

TTGGACCCGT CGTGATGTGG ACCTTAATAA TACGATTACC CCAAATAGAA TTACACAATT 2220

ACCATTGGTA AAGGCATCTG CACCTGTTTC GGGTACTACG GTCTTAAAAG GTCCAGGATT 2280

TACAGGAGGG GGTATACTCC GAAGAACAAC TAATGGCACA TTTGGAACGT TAAGAGTAAC 2340

GGTTAATTCA CCATTAACAC AACAATATCG CCTAAGAGTT CGTTTTGCCT CAACAGGAAA 2400

TTTCAGTATA AGGGTACTCC GTGGAGGGGT TTCTATCGGT GATGTTAGAT TAGGGAGCAC 2460

AATGAACAGA GGGCAGGAAC TAACTTACGA ATCCTTTTTC ACAAGAGAGT TTACTACTAC 2520

TGGTCCGTTC AATCCGCCTT TTACATTTAC ACAAGCTCAA GAGATTCTAA CAGTGAATGC 2580

AGAAGGTGTT AGCACCGGTG GTGAATATTA TATAGATAGA ATTGAAATTG TCCCTGTGAA 2640

TCCGGCACGA GAAGCGGAAG AGGATTTAGA AGCGGCGAAG AAAGCGGTGG CGAGCTTGTT 2700

TACACGTACA AGGGACGGAT TACAGGTAAA TGTGACAGAT TATCAAGTGG ACCAAGCGGC 2760

AAATTTAGTG TCATGCTTAT CCGATGAACA ATATGGGCAT GACAAAAAGA TGTTATTGGA 2820

AGCGGTAAGA GCGGCAAAAC GCCTCAGCCG CGAACGCAAC TTACTTCAAG ATCCAGATTT 2880

TAATACAATC AATAGTACAG AAGAGAATGG CTGGAAGGCA AGTAACGGTG TTACTATTAG 2940

CGAGGGCGGT CCATTCTTTA AAGGTCGTGC ACTTCAGTTA GCAAGCGCAA GAGAAAATTA 3000

TCCAACATAC ATTTATCAAA AAGTAGATGC ATCGGTGTTA AAGCCTTATA CACGCTATAG 3060

ACTAGATGGA TTTGTGAAGA GTAGTCAAGA TTTAGAAATT GATCTCATCC ACCATCATAA 3120

AGTCCATCTT GTAAAAAATG TACCAGATAA TTTAGTATCT GATACTTACT CAGATGGTTC 3180

TTGCAGCGGA ATCAACCGTT GTGATGAACA GCATCAGGTA GATATGCAGC TAGATGCGGA 3240

GCATCATCCA ATGGATTGCT GTGAAGCGGC TCAAACACAT GAGTTTTCTT CCTATATTAA 3300

TACAGGGGAT CTAAATGCAA GTGTAGATCA GGGCATTTGG GTTGTATTAA AAGTTCGAAC 3360

AACAGATGGG TATGCGACGT TAGGAAATCT TGAATTGGTA GAGGTTGGGC CATTATCGGG 3420

TGAATCTCTA GAACGGGAAC AAAGAGATAA TGCGAAATGG AATGCAGAGC TAGGAAGAAA 3480

ACGTGCAGAA ATAGATCGTG TGTATTTAGC TGCGAAACAA GCAATTAATC ATCTGTTTGT 3540

AGACTATCAA GATCAACAAT TAAATCCAGA AATTGGGCTA GCAGAAATTA ATGAAGCTTC 3600

AAATCTTGTA GAGTCAATTT CGGGTGTATA TAGTGATACA CTATTACAGA TTCCTGGGAT 3660

TAACTACGAA ATTTACACAG AGTTATCCGA TCGCTTACAA CAAGCATCGT ATCTGTATAC 3720

GTCTAGAAAT GCGGTGCAAA ATGGAGACTT TAACAGTGGT CTAGATAGTT GGAATACAAC 3780

TATGGATGCA TCGGTTCAGC AAGATGGCAA TATGCATTTC TTAGTTCTTT CGCATTGGGA 3840

TGCACAAGTT TCCCAACAAT TGAGAGTAAA TCCGAATTGT AAGTATGTCT TACGTGTGAC 3900

AGCAAGAAAA GTAGGAGGCG GAGATGGATA CGTCACAATC CGAGATGGCG CTCATCACCA 3960

AGAAACTCTT ACATTTAATG CATGTGACTA CGATGTAAAT GGTACGTATG TCAATGACAA 4020

TTCGTATATA ACAGAAGAAG TGGTATTCTA CCCAGAGACA AAACATATGT GGGTAGAGGT 4080

GAGTGAATCC GAAGGTTCAT TCTATATAGA CAGTATTGAG TTTATTGAAA CACAAGAGTA 4140

GAAGAGGGGG ATCCTAACGT ATAGCAACTA TGAGAGGATA CTCCGTACAA ACAAAGATTA 4200

AAAAAAGGTA AAATGAATAG AACCCCCTAC TGGTAGAAGG ACCGATAGGG GGTTCTTACA 4260

TGAAAAAATG TAGCTGTTTA CTAAGGTGTA TAAAAAACAG CATATCTGAT AGAAAAAAGT 4320

GAGTACCTTA TAAAGAAAGA ATTC 4344

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1157 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His
1 5 10 15

Cys Gly Cys Pro Ser Asp Asp Asp Val Arg Tyr Pro Leu Ala Ser Asp
20 25 30

Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met
35 40 45

Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile
50 55 60

Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile
65 70 75 80

Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr
85 90 95

Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp
100 105 110

Glu Ala Phe Met Arg Gln Val Glu Glu Leu Val Asn Gln Gln Ile Thr
115 120 125

Glu Phe Ala Arg Asn Gln Ala Leu Ala Arg Leu Gln Gly Leu Gly Asp
130 135 140

Ser Phe Asn Val Tyr Gln Arg Ser Leu Gln Asn Trp Leu Ala Asp Arg
145 150 155 160

Asn Asp Thr Arg Asn Leu Ser Val Val Arg Ala Gln Phe Ile Ala Leu
165 170 175

Asp Leu Asp Phe Val Asn Ala Ile Pro Leu Phe Ala Val Asn Gly Gln
180 185 190

Gln Val Pro Leu Leu Ser Val Tyr Ala Gln Ala Val Asn Leu His Leu
195 200 205

Leu Leu Leu Lys Asp Ala Ser Leu Phe Gly Glu Gly Trp Gly Phe Thr
210 215 220

```
Gln Gly Glu Ile Ser Thr Tyr Tyr Asp Arg Gln Leu Glu Leu Thr Ala
225                 230                 235                 240

Lys Tyr Thr Asn Tyr Cys Glu Thr Trp Tyr Asn Thr Gly Leu Asp Arg
                245                 250                 255

Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Arg Tyr His Gln Phe Arg
            260                 265                 270

Arg Glu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr
        275                 280                 285

Tyr Asp Val Arg Leu Tyr Pro Thr Gly Ser Asn Pro Gln Leu Thr Arg
    290                 295                 300

Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Pro Pro Ala Asn Val Gly
305                 310                 315                 320

Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser Glu Leu
                325                 330                 335

Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu Asn Ser
            340                 345                 350

Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp
        355                 360                 365

Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala
    370                 375                 380

Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile
385                 390                 395                 400

Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp
                405                 410                 415

Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe
            420                 425                 430

Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly
        435                 440                 445

Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser
    450                 455                 460

Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe
465                 470                 475                 480

Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr
                485                 490                 495

Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro
            500                 505                 510

Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser
        515                 520                 525

Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu
    530                 535                 540

Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn
545                 550                 555                 560
```

Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr
565 570 575

Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Val Ser Ile Gly Asp
580 585 590

Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu
595 600 605

Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro
610 615 620

Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly
625 630 635 640

Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro
645 650 655

Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys
660 665 670

Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn
675 680 685

Val Thr Asp Tyr Gln Val Asp Gln Ala Ala Asn Leu Val Ser Cys Leu
690 695 700

Ser Asp Glu Gln Tyr Gly His Asp Lys Lys Met Leu Leu Glu Ala Val
705 710 715 720

Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro
725 730 735

Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser
740 745 750

Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly Arg Ala
755 760 765

Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln
770 775 780

Lys Val Asp Ala Ser Val Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp
785 790 795 800

Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His
805 810 815

His Lys Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Ser Asp
820 825 830

Thr Tyr Ser Asp Gly Ser Cys Ser Gly Ile Asn Arg Cys Asp Glu Gln
835 840 845

His Gln Val Asp Met Gln Leu Asp Ala Glu His His Pro Met Asp Cys
850 855 860

Cys Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly
865 870 875 880

Asp Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val
885 890 895

```
Arg Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu
            900                 905                 910

Val Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn
        915                 920                 925

Ala Lys Trp Asn Ala Glu Leu Gly Arg Lys Arg Ala Glu Ile Asp Arg
    930                 935                 940

Val Tyr Leu Ala Ala Lys Gln Ala Ile Asn His Leu Phe Val Asp Tyr
945                 950                 955                 960

Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly Leu Ala Glu Ile Asn Glu
                965                 970                 975

Ala Ser Asn Leu Val Glu Ser Ile Ser Gly Val Tyr Ser Asp Thr Leu
            980                 985                 990

Leu Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asp
        995                 1000                1005

Arg Leu Gln Gln Ala Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val Gln
    1010                1015                1020

Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Thr Thr Met Asp
1025                1030                1035                1040

Ala Ser Val Gln Gln Asp Gly Asn Met His Phe Leu Val Leu Ser His
                1045                1050                1055

Trp Asp Ala Gln Val Ser Gln Gln Leu Arg Val Asn Pro Asn Cys Lys
            1060                1065                1070

Tyr Val Leu Arg Val Thr Ala Arg Lys Val Gly Gly Gly Asp Gly Tyr
        1075                1080                1085

Val Thr Ile Arg Asp Gly Ala His His Gln Glu Thr Leu Thr Phe Asn
    1090                1095                1100

Ala Cys Asp Tyr Asp Val Asn Gly Thr Tyr Val Asn Asp Asn Ser Tyr
1105                1110                1115                1120

Ile Thr Glu Glu Val Val Phe Tyr Pro Glu Thr Lys His Met Trp Val
                1125                1130                1135

Glu Val Ser Glu Ser Glu Gly Ser Phe Tyr Ile Asp Ser Ile Glu Phe
            1140                1145                1150

Ile Glu Thr Gln Glu
        1155
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1897 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION:13..1890
(D) OTHER INFORMATION:/note= "coding sequence"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGTACCAAAA CCATGGCTGA CTACCTGCAG ATGACCGACG AGGACTACAC CGACAGCTAC 60

ATCAACCCCA GCCTGAGCAT CAGCGGTCGC GACGCCGTGC AGACCGCTCT GACCGTGGTG 120

GGTCGCATCC TGGGTGCCCT GGGCGTGCCC TTCAGCGGTC AGATCGTGAG CTTCTACCAG 180

TTCCTGCTGA ACACCCTGTG GCCAGTGAAC GACACCGCCA TCTGGGAAGC TTTCATGCGC 240

CAGGTGGAGG AGCTGGTGAA CCAGCAGATC ACCGAGTTCG CTCGCAACCA GGCCCTGGCT 300

CGCCTGCAGG GCCTGGGCGA CAGCTTCAAC GTGTACCAGC GCAGCCTGCA GAACTGGCTG 360

GCCGACCGCA ACGACACCAA GAACCTGAGC GTGGTGAGGG CCCAGTTCAT CGCCCTGGAC 420

CTGGACTTCG TGAACGCCAT CCCCCTGTTC GCCGTGAACG GCCAGCAGGT GCCCCTGCTG 480

AGCGTGTACG CCCAGGCCGT GAACCTGCAC CTGCTGCTGC TGAAGGATGC ATCCCTGTTC 540

GGCGAGGGCT GGGGCTTCAC CCAGGGCGAG ATCAGCACCT ACTACGACCG CCAGCTCGAG 600

CTGACCGCCA AGTACACCAA CTACTGCGAG ACCTGGTACA ACACCGGTCT GGACCGCCTG 660

AGGGGCACCA ACACCGAGAG CTGGCTGCGC TACCACCAGT TCCGCAGGGA GATGACCCTG 720

GTGGTGCTGG ACGTGGTGGC CCTGTTCCCC TACTACGACG TGCGCCTGTA CCCCACCGGC 780

AGCAACCCCC AGCTGACACG TGAGGTGTAC ACCGACCCCA TCGTGTTCAA CCCACCAGCC 840

AACGTGGGCC TGTGCCGCAG GTGGGGCACC AACCCCTACA ACACCTTCAG CGAGCTGGAG 900

AACGCCTTCA TCAGGCCACC CCACCTGTTC GACCGCCTGA ACAGCCTGAC CATCAGCAGC 960

AATCGATTCC CCGTGAGCAG CAACTTCATG GACTACTGGA GCGGTCACAC CCTGCGCAGG 1020

AGCTACCTGA ACGACAGCGC CGTGCAGGAG GACAGCTACG GCCTGATCAC CACCACCAGG 1080

GCCACCATCA ACCCAGGCGT GGACGGCACC AACCGCATCG AGAGCACCGC TGTGGACTTC 1140

CGCAGCGCTC TGATCGGCAT CTACGGCGTG AACAGGGCCA GCTTCGTGCC AGGTGGCCTG 1200

TTCAACGGCA CCACCAGCCC AGCCAACGGT GGCTGCCGAG ATCTGTACGA CACCAACGAC 1260

GAGCTGCCAC CCGACGAGAG CACCGGCAGC AGCACCCACC GCCTGAGCCA CGTCACCTTC 1320

TTCAGCTTCC AGACCAACCA GGCTGGCAGC ATCGCCAACG CTGGCAGCGT GCCCACCTAC 1380

GTGTGGACCA GGAGGGACGT GGACCTGAAC AACACCATCA CCCCCAACCG CATCACCCAG 1440

CTGCCCCTGG TGAAGGCCAG CGCTCCCGTG AGCGGCACCA CCGTGCTGAA GGGTCCAGGC 1500

TTCACCGGTG GCGGTATACT GCGCAGGACC ACCAACGGCA CCTTCGGCAC CCTGCGCGTG 1560

ACCGTGAATT CCCCACTGAC CCAGCAGTAC CGCCTGCGCG TGCGCTTCGC CAGCACCGGC 1620

AACTTCAGCA TCCGCGTGCT GAGGGGTGGC GTGAGCATCG GCGACGTGCG CCTGGGCAGC 1680

ACCATGAACA GGGGCCAGGA GCTGACCTAC GAGAGCTTCT TCACCCGCGA GTTCACCACC 1740

ACCGGTCCCT TCAACCCACC CTTCACCTTC ACCCAGGCCC AGGAGATCCT GACCGTGAAC 1800

GCCGAGGGCG TGAGCACCGG TGGCGAGTAC TACATCGACC GCATCGAGAT CGTGCCCGTG 1860

AACCCAGCTC GCGAGGCCGA GGAGGACTGA GGCTAGC 1897

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 625 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met Ala Asp Tyr Leu Gln Met Thr Asp Glu Asp Tyr Thr Asp Ser Tyr
1 5 10 15

Ile Asn Pro Ser Leu Ser Ile Ser Gly Arg Asp Ala Val Gln Thr Ala
20 25 30

Leu Thr Val Val Gly Arg Ile Leu Gly Ala Leu Gly Val Pro Phe Ser
35 40 45

Gly Gln Ile Val Ser Phe Tyr Gln Phe Leu Leu Asn Thr Leu Trp Pro
50 55 60

Val Asn Asp Thr Ala Ile Trp Glu Ala Phe Met Arg Gln Val Glu Glu
65 70 75 80

Leu Val Asn Gln Gln Ile Thr Glu Phe Ala Arg Asn Gln Ala Leu Ala
85 90 95

Arg Leu Gln Gly Leu Gly Asp Ser Phe Asn Val Tyr Gln Arg Ser Leu
100 105 110

Gln Asn Trp Leu Ala Asp Arg Asn Asp Thr Lys Asn Leu Ser Val Val
115 120 125

Arg Ala Gln Phe Ile Ala Leu Asp Leu Asp Phe Val Asn Ala Ile Pro
130 135 140

Leu Phe Ala Val Asn Gly Gln Gln Val Pro Leu Leu Ser Val Tyr Ala
145 150 155 160

Gln Ala Val Asn Leu His Leu Leu Leu Leu Lys Asp Ala Ser Leu Phe
165 170 175

Gly Glu Gly Trp Gly Phe Thr Gln Gly Glu Ile Ser Thr Tyr Tyr Asp
180 185 190

Arg Gln Leu Glu Leu Thr Ala Lys Tyr Thr Asn Tyr Cys Glu Thr Trp
195 200 205

```
Tyr Asn Thr Gly Leu Asp Arg Leu Arg Gly Thr Asn Thr Glu Ser Trp
    210                 215                 220

Leu Arg Tyr His Gln Phe Arg Arg Glu Met Thr Leu Val Val Leu Asp
225                 230                 235                 240

Val Val Ala Leu Phe Pro Tyr Tyr Asp Val Arg Leu Tyr Pro Thr Gly
                245                 250                 255

Ser Asn Pro Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Ile Val Phe
            260                 265                 270

Asn Pro Pro Ala Asn Val Gly Leu Cys Arg Arg Trp Gly Thr Asn Pro
        275                 280                 285

Tyr Asn Thr Phe Ser Glu Leu Glu Asn Ala Phe Ile Arg Pro Pro His
    290                 295                 300

Leu Phe Asp Arg Leu Asn Ser Leu Thr Ile Ser Ser Asn Arg Phe Pro
305                 310                 315                 320

Val Ser Ser Asn Phe Met Asp Tyr Trp Ser Gly His Thr Leu Arg Arg
                325                 330                 335

Ser Tyr Leu Asn Asp Ser Ala Val Gln Glu Asp Ser Tyr Gly Leu Ile
            340                 345                 350

Thr Thr Thr Arg Ala Thr Ile Asn Pro Gly Val Asp Gly Thr Asn Arg
        355                 360                 365

Ile Glu Ser Thr Ala Val Asp Phe Arg Ser Ala Leu Ile Gly Ile Tyr
    370                 375                 380

Gly Val Asn Arg Ala Ser Phe Val Pro Gly Gly Leu Phe Asn Gly Thr
385                 390                 395                 400

Thr Ser Pro Ala Asn Gly Gly Cys Arg Asp Leu Tyr Asp Thr Asn Asp
                405                 410                 415

Glu Leu Pro Pro Asp Glu Ser Thr Gly Ser Ser Thr His Arg Leu Ser
            420                 425                 430

His Val Thr Phe Phe Ser Phe Gln Thr Asn Gln Ala Gly Ser Ile Ala
        435                 440                 445

Asn Ala Gly Ser Val Pro Thr Tyr Val Trp Thr Arg Arg Asp Val Asp
    450                 455                 460

Leu Asn Asn Thr Ile Thr Pro Asn Arg Ile Thr Gln Leu Pro Leu Val
465                 470                 475                 480

Lys Ala Ser Ala Pro Val Ser Gly Thr Thr Val Leu Lys Gly Pro Gly
                485                 490                 495

Phe Thr Gly Gly Gly Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly
            500                 505                 510

Thr Leu Arg Val Thr Val Asn Ser Pro Leu Thr Gln Gln Tyr Arg Leu
        515                 520                 525

Arg Val Arg Phe Ala Ser Thr Gly Asn Phe Ser Ile Arg Val Leu Arg
    530                 535                 540
```

```
Gly Gly Val Ser Ile Gly Asp Val Arg Leu Gly Ser Thr Met Asn Arg
545                 550                 555                 560

Gly Gln Glu Leu Thr Tyr Glu Ser Phe Phe Thr Arg Glu Phe Thr Thr
                565                 570                 575

Thr Gly Pro Phe Asn Pro Pro Phe Thr Phe Thr Gln Ala Gln Glu Ile
            580                 585                 590

Leu Thr Val Asn Ala Glu Gly Val Ser Thr Gly Gly Glu Tyr Tyr Ile
        595                 600                 605

Asp Arg Ile Glu Ile Val Pro Val Asn Pro Ala Arg Glu Ala Glu Glu
    610                 615                 620

Asp
625
```

Claims

1. A modified Cry9C protein with an improved toxicity to an insect species, comprising the amino acid sequence of SEQ ID No. 2 or an insecticidally-effective fragment thereof, wherein at least one amino acid in one of the following regions in SEQ ID No. 2 is replaced by another amino acid: 313-334, 358-369, 418-425, 480-492.

2. A modified Cry9C protein with an improved toxicity to an insect species, comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid 658, wherein at least one of the amino acids at the following positions in SEQ ID No. 2 have been replaced by another amino acid: 313, 316, 317, 318, 319, 321, 323, 325, 329, 330, 368, 369, 418, 420, 421, 422, 480, 481, 483, 484, 485, 487, 488, 490 and 491.

3. The modified Cry9C protein of claim 1 wherein said at least one amino acid position is position 316, 317, 319, 321, 329, 330, 369, 422, or 488 in SEQ ID No. 2.

4. The modified Cry9C protein of claim 1 with improved toxicity to *Ostrinia nubilalis*, comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein at least the amino acids at positions 364 and 488 in SEQ ID No. 2 are replaced by other amino acids.

5. The modified Cry9C protein of claim 1 with improved toxicity to *Heliothis virescens*, comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein the amino acid at position 321 or position 329 in SEQ ID No 2, is replaced by another amino acid.

6. The modified Cry9C protein of claim 1 with improved toxicity to *Diatraea grandiosella*, comprising the amino acid sequence of SEQ ID No. 2 from amino acid position 1 or 44 to amino acid position 658, wherein the amino acid at any or all of

amino acid positions 316, 317, 319, 321, 330, 369, or 422 in SEQ ID No. 2 is replaced by another amino acid.

7. The modified Cry9C protein of any one of claims 1 to 6 wherein the arginine at position 164 in SEQ ID No. 2 is replaced by another amino acid.

8. The modified Cry9C protein of any one of claims 1 to 6 wherein said at least one amino acid position is replaced by alanine.

9. A DNA sequence encoding the protein of any one of claims 1 to 6.

10. A DNA sequence encoding the protein of claim 7 or 8.

11. A plant, comprising the DNA of claim 9 or 10.

12. A seed, comprising the DNA of claim 9 or 10.

13. The plant of claim 11 which is selected from the group consisting of: corn, cotton, rice, oilseed rape, cauliflower, broccoli, soybean, tomato, tobacco, potato, eggplant, beet, oat, pepper, gladiolus, dahlia, chrysanthemum, sorghum, and garden peas.

14. A method for controlling insects feeding on a plant, comprising expressing the protein of any one of claims 1 to 6 in a plant.

15. A method for controlling insects feeding on a plant, comprising growing the plant of Claim 11.

16. A method of obtaining a seed comprising the DNA of Claim 9 or 10 comprising inserting said DNA into the genome of a plant and harvesting the seed from said plant.